What is claimed is:

1. A method for collecting a document from a network, comprising:

collecting documents equal to or larger, in number, than a predetermined value from inside a community through the network based on a reference of the document; and

collecting documents from inside and outside the community based on the reference of collected documents after collecting the documents equal to or larger in number than the predetermined value from inside the community.

2. The method according to claim 1, further comprising:

computing a significance level indicating a level of significance of the collected document according to the reference of the collected document, and information about a position of the collected document in the network; and

determining a document to be collected based on the reference and the significance level.

3. The method according to claim 2, wherein

said document to be collected is determined separately for inside the community and for outside the community.

4. The method according to claim 3, further comprising:

presenting a result of retrieving the collected documents separately for inside the community and outside the community.

5. The method according to claim 2, further comprising:

determining whether or not the document is in the community according to information indicating the position of the document in the network.

6. A method for collecting a document from a network, comprising:

providing a positive sample document group which is a document group relating to a field, and a negative sample document group which is a document group less related to the field;

determining a document which is to be collected and is related to the field based on a reference to the positive sample document group and

the negative sample document group; and

collecting the document to be collected from the network.

7. The method according to claim 6, further comprising:

computing a reference score indicating a level at which a document is referenced only by a document in the positive sample document group based on the reference; and

determining a document having a high reference score as the document to be collected.

8. The method according to claim 6, wherein

computing a co-reference score indicaiting a level at which a document is referenced together with a document in the positive sample document group for a document referenced by a collected document referring to a document in the positive sample document group based on the reference; and

determining a document having a high co-reference score as the document to be collected.

9. The method according to claim 6, wherein

said negative sample document group is a union

of document groups relating to a plurality of fields.

10. The method according to claim 1, further comprising:

summarizing said collected document group based on a referencing expression used in the collected document group.

11. The method according to claim 1, further comprising:

assigning a keyword to the collected document based on a referencing expression used in the collected document.

12. The method according to claim 1, further comprising:

not assigning a keyword based on the referring expression when the referencing expression is used regardless of a content of a referenced document.

13. The method according to claim 11, further comprising:

counting a number of different documents referenced using the referencing expression; and

not assigning the keyword based on the referencing expression when the number of different documents is equal to or larger than a predetermined value.

14. The method according to claim 11, further comprising:

counting a reference frequency at which each collected document is referenced by the referencing expression when the number of different documents is smaller than a predetermined value; and

determining whether or not the referencing expression is assigned as the keyword based on the number of different documents and the reference frequency.

15. The method according to claim 11, further comprising:

combining the keyword based on the referencing expression with a keyword extracted from text of the collected document, and a keyword extracted from information indicating a position in the network about the collected document.

16. A method for retrieving a document from a

terminal belonging to a community in a network, comprising:

transmitting information for retrieval of the document to a server; and

receiving the document retrieved separately from inside and outside the community according to the information for retrieval together with information indicating a significance level for the community.

17. A document collection apparatus collecting a document from a network, comprising:

a next prospect determination unit determining a prospect to be collected next based on a reference of a collected document;

a community determination unit determining whether or not the prospect is in a community in the network according to information indicating a position in the network of the prospect; and

a document collection unit collecting the prospect from the network, wherein

said document collection unit collects the prospect from inside and outside the community after collecting documents larger in number than a predetermined value from inside the community.

18. A document collection apparatus collecting a document from a network, comprising:

a next prospect determination unit determining a prospect to be collected next based on a reference between a positive sample document group which is a document group related to a field and a negative sample document group which is a document group less related the field; and

a document collection unit collecting the prospect from the network.

19. A computer-readable recording medium recording a program used to direct a computer to control collection of a document from a network, comprising:

collecting documents equal to or larger, in number, than a predetermined value from a community through the network based on a reference of the document; and

collecting documents from inside and outside the community based on the reference of collected documents after collecting the documents equal to or larger, in number, than the predetermined value from inside the community.

20. A computer-readable recording medium recording a program used to direct a computer to control collection of a document from a network, comprising:

providing a positive sample document group which is a document group relating to a field, and a negative sample document group which is a document group less related to the field;

determining a document to be collected relating to the field based on a reference to the positive sample document group and the negative sample document group; and

collecting the document to be collect from the network.

21. A computer data signal embodied on a carrier expressing a program used to direct a computer to control collection of a document from a network, said program allowing the computer to perform the process comprising:

collecting documents equal to or larger than, in number, a predetermined value from inside a community in the network based on a reference of the document; and

collecting documents from inside and outside the community based on the reference of collected documents after collecting documents equal to or larger, in number, than the predetermined value from the community.